Computer Life
ISSN: 1819-4818 | Vol. 14, No. 1, 2026

Adaptive AI Tutors Scale Personalized STEM Education Across Diverse Learner Populations

Zhenyu Luo *, Haotian Ren, and Matthew Carter

Department of Computer Science, University of Central Florida, USA
*** Corresponding author Email:** luo8372891@gmail.com

Abstract: The persistent challenge of delivering individualized instruction across heterogeneous learner populations in science, technology, engineering, and mathematics (STEM) contexts has driven significant interest in artificial intelligence (AI)-powered adaptive tutoring systems. This paper examines how adaptive AI tutors can be effectively scaled to meet the cognitive, motivational, and demographic diversity of students engaged in STEM learning. Drawing on a mixed-methods research design that integrates machine learning-based knowledge tracing, reinforcement learning (RL)-driven content sequencing, and learner profile modeling, this study evaluates an adaptive intelligent tutoring system (ITS) deployed across three secondary and tertiary STEM courses involving a combined cohort of 847 students. Quantitative findings indicate statistically significant improvements in learning gains, task completion rates, and knowledge retention among students who engaged with the adaptive system compared to those in conventional instructional settings. Qualitative data further reveals that learners from under-resourced backgrounds and high-achieving students both reported increased perceived relevance and self-regulatory engagement when the system dynamically adjusted content difficulty and scaffolding level. The study also identifies equity-related concerns regarding data sparsity for underrepresented learner groups and proposes mitigation strategies anchored in fairness-aware machine learning. These findings contribute to an emerging understanding of how scalable adaptive AI tutoring can bridge achievement gaps in STEM education without sacrificing instructional depth or pedagogical coherence.

Keywords: Adaptive tutoring systems; Intelligent tutoring systems; Personalized STEM learning; Knowledge tracing; Learner diversity; Educational AI; Reinforcement learning.

1. Introduction

The global imperative to expand access to high-quality STEM education has intensified in recent decades, driven by rapid technological transformation, shifting labor market demands, and mounting recognition of persistent achievement disparities across gender, socioeconomic, linguistic, and cultural dimensions. Traditional classroom instruction, constrained by fixed pacing, standardized curricula, and limited teacher-to-student ratios, has long struggled to accommodate the substantial variability in prior knowledge, learning preferences, and motivational profiles that characterizes any realistic student population. This structural tension between scalability and personalization has become one of the defining challenges facing STEM education systems worldwide [1]. The aspiration to provide every student with the quality of individualized guidance historically reserved for those with access to private tutors or elite institutional resources is no longer merely a pedagogical ideal; it has become a matter of educational equity and economic urgency [2].

Artificial intelligence offers a compelling response to this challenge. Over the past decade, AI-powered adaptive tutoring technologies have evolved from rule-based expert systems into sophisticated platforms capable of modeling individual learner states in real time, dynamically adjusting instructional content and difficulty, and providing formative feedback that responds to the specific nature of a students misconceptions. Unlike static e-learning environments, adaptive ITS integrate multiple computational layers including deep knowledge tracing (DKT), pedagogical sequencing algorithms, and natural language processing (NLP) to approximate the one-on-one tutorial interaction that research has long identified as educationally optimal [3]. The promise of such systems lies not merely in their personalization capabilities, but in their potential to be deployed at population scale, reaching thousands of learners simultaneously without proportional increases in instructional cost [4].

The relevance of adaptive AI tutors to STEM education is particularly acute. STEM subjects are characterized by hierarchical knowledge structures, sequential competency dependencies, and a requirement for procedural fluency that makes gap-free foundational understanding essential for subsequent progress. A student who fails to consolidate algebraic manipulation will encounter compounding difficulties as they advance to calculus or statistics. Conventional instruction often lacks the diagnostic sensitivity to detect and address such gaps before they become entrenched [5]. Adaptive ITS, by contrast, can track mastery at the level of individual knowledge components (KCs), triggering targeted remediation precisely when and where learning breakdowns occur. This capability is particularly valuable for diverse learner populations in which prior knowledge varies enormously and the moment of conceptual difficulty is highly individual [6].

The diversity dimension of this problem has received increasing scholarly attention. Learner populations in contemporary STEM programs encompass students from different cultural backgrounds, those with varying degrees of prior STEM exposure, learners with disabilities or specific learning needs, and students navigating significant socioeconomic constraints on study time and resource access [7]. A central concern is whether adaptive AI systems, trained predominantly on data generated by majority populations, can equitably serve learners whose interaction patterns, terminology use, or problem-solving approaches diverge from the training distribution. This concern intersects with broader debates about algorithmic fairness, data ethics, and the potential for AI systems to reproduce or amplify existing educational inequalities [8]. The question of whose learning experience is optimized — and whose is inadvertently neglected — by intelligent tutoring technology is not merely technical; it is fundamentally ethical and political.

Despite considerable progress in building effective adaptive tutoring technologies, the challenge of scaling these systems across truly diverse populations remains underexplored. Much existing research has been conducted in controlled, resource-rich settings with relatively homogeneous student cohorts [9]. Questions remain about how adaptive AI tutors perform when deployed in under-resourced schools, when learner interaction data is sparse or noisy, or when students bring cultural and linguistic contexts not represented in the systems training corpus. Addressing these questions requires not only technical innovation in machine learning and learner modeling, but also careful attention to participatory design principles that incorporate teacher and student perspectives into system development [10]. The intersection of pedagogical theory, equity scholarship, and computational innovation is therefore the essential context in which this study is situated.

This study contributes to this conversation by reporting on the design, implementation, and evaluation of an adaptive AI tutoring system deployed across diverse secondary and tertiary STEM learning contexts. The system integrates DKT for real-time learner state estimation, an RL-based curriculum sequencer for adaptive content selection, and an equity-sensitive personalization module designed to ensure that minority learner groups receive instructionally appropriate rather than data-deficient experiences.

2. Literature Review

The origins of adaptive ITS can be traced to early cognitive tutors developed in the 1980s and 1990s, which modeled student knowledge as probability distributions over discrete skill components and adjusted problem selection accordingly. The field has since undergone profound transformation in the era of deep learning and large-scale educational data [11]. Contemporary research on ITS draws on insights from cognitive science, learning analytics, educational psychology, and computer science, yielding systems of considerably greater sophistication and applicability than their predecessors. The convergence of these disciplines has produced a research agenda defined by a commitment to evidence-based design, empirical validation at scale, and increasing sensitivity to the equity implications of algorithmic decision-making in educational settings [12].

A foundational concern in adaptive tutoring system design is the accuracy of the underlying learner model. Knowledge tracing — the task of estimating a students mastery of individual skills over time based on observed response sequences — has long been dominated by Bayesian knowledge tracing (BKT), which treats each skill as a latent binary variable evolving under Markov dynamics. While BKT captures certain essential features of skill acquisition, it makes strong independence assumptions that limit its ability to represent the complex interdependencies among knowledge components characteristic of STEM subjects [13]. The introduction of deep learning approaches to knowledge tracing, embodied in the DKT framework using recurrent neural networks (RNNs), represented a significant paradigm shift. Specifically, DKT employs long short-term memory (LSTM) units to maintain a continuous hidden state that evolves with each observed student interaction, enabling the model to capture far richer dependency structures than BKT permits. This architectural insight — that student knowledge can be represented as a continuously evolving latent vector rather than a discrete binary state — substantially improved predictive accuracy across multiple real-world educational datasets [14]. Subsequent work extended the DKT framework using attention mechanisms and dynamic memory networks to better represent concept-level relational structures in complex STEM domains [15].

The scalability of personalized STEM tutoring has been a central preoccupation of recent literature. Research on adaptive learning in higher education has documented consistent but often modest improvements in learning outcomes when adaptive systems replace or supplement traditional instruction [16]. The heterogeneity of findings across studies reflects the sensitivity of adaptive systems to implementation context, including the degree of teacher involvement, the alignment between system content and course objectives, and the motivation and digital literacy of the student population. Systematic reviews have emphasized that adaptive learning technology is most effective when embedded in pedagogically coherent course designs, where instructors actively use system-generated learning analytics to inform their own instructional decisions rather than treating the technology as a standalone intervention. This finding underscores the importance of framing adaptive AI tutoring as a teacher-enhancement tool rather than a teacher-replacement strategy, a distinction with important implications for both system design and institutional adoption.

In STEM-specific contexts, the evidence base for adaptive tutoring is particularly robust in mathematics. Systems built around cognitive tutor principles have accumulated substantial empirical evidence of learning benefits across diverse K–12 populations, including students from low-income backgrounds [17]. Studies examining these systems have highlighted the importance of intelligent error feedback, worked examples adaptively sequenced to the learners current proficiency, and the capacity for the system to identify not merely whether a student answered correctly but why they likely erred. Physics and chemistry tutoring systems have followed similar design trajectories, with platforms incorporating natural language dialogue and simulation-based problem solving to support conceptual understanding alongside procedural skill. This multimodal approach to adaptive instruction is increasingly recognized as essential for addressing the full complexity of STEM learning, which spans procedural automaticity, conceptual understanding, and the capacity for novel problem transfer across disciplinary contexts.

The equity dimensions of adaptive learning have received growing scholarly attention in recent years. Research on teacher and student needs for AI-enhanced classrooms has emphasized the need for adaptive tutoring to support teacher agency rather than displacing it, with the complementarity between human and algorithmic judgment identified as essential for serving students with complex needs [18]. From a data justice perspective, scholars have argued that explainable AI in education is not merely a technical design preference but an ethical imperative: learners and educators deserve transparency about the algorithmic processes shaping instructional decisions [19]. Studies examining differential system performance across demographic groups have documented concerning patterns in knowledge tracing models trained on unbalanced data, where underrepresented learner groups may receive less accurate or less adaptive responses than their better-represented peers [20]. Addressing these disparities requires both technical innovation and institutional commitment to ongoing equity auditing of deployed adaptive systems.

The role of RL in adaptive curriculum sequencing has emerged as a particularly active research frontier. Unlike supervised learning approaches that optimize moment-to-moment prediction accuracy, RL-based pedagogical agents optimize long-term learning trajectories by selecting instructional actions — such as the choice of problem type, difficulty level, or scaffolding format — to maximize cumulative learning gain [21]. Early implementations demonstrated proof of concept in constrained laboratory settings, but recent work has extended RL-based tutoring to realistic online learning environments with thousands of concurrent users, achieving improvements in knowledge retention compared to static curricula. The integration of RL with learner engagement modeling — using signals such as response latency, error patterns, and hint request frequency — has further enhanced the responsiveness of adaptive systems to within-session fluctuations in student motivation and cognitive load [22].

NLP has become an increasingly central component of adaptive tutoring systems, enabling richer forms of formative feedback and student-tutor dialogue. NLP-enabled ITS can parse free-text student responses to identify the presence or absence of specific conceptual elements, enabling the system to deliver targeted feedback addressing the precise nature of a students misconception rather than simply flagging an answer as incorrect [23]. Recent advances in generative large language models (LLMs) have opened new possibilities for generating coherent, contextually appropriate instructional dialogue at scale, though questions remain about the reliability, interpretability, and cultural appropriateness of LLM-generated educational content across diverse learner populations [24]. The intersection of NLP, learner modeling, and pedagogical theory represents one of the most technically complex and educationally consequential frontiers in contemporary adaptive tutoring research.

Despite the accumulation of evidence documenting the benefits of adaptive AI tutors in STEM education, significant gaps remain in the literature. Most studies have been conducted in North American or European educational contexts, limiting their generalizability to diverse global settings [25]. Experimental evaluations have frequently excluded learners with disabilities or those requiring specialized instructional accommodations, leaving important equity questions unresolved. Additionally, the temporal validity of adaptive tutoring effects — whether gains achieved with AI tutor assistance persist and transfer to new learning contexts without the system — remains underexplored [26]. The present study seeks to address several of these gaps by evaluating adaptive AI tutoring across culturally diverse cohorts, by including explicit analysis of differential system performance across learner demographic subgroups, and by examining the mediating role of teacher engagement in shaping student outcomes.

3. Methodology

3.1. System Architecture and Learner Modeling

The adaptive AI tutoring system developed for this study was organized around three core computational components: a learner modeling module based on DKT, a curriculum sequencing engine driven by RL, and a personalization layer that integrates learner demographic and contextual metadata to ensure equitable instructional delivery. These components were implemented as modular microservices communicating through a RESTful application programming interface (API), enabling the system to scale horizontally across concurrent user sessions without performance degradation. The modular design also facilitated component-level ablation studies, allowing researchers to isolate the contribution of each subsystem to overall learning outcomes.

The learner modeling module implements a bidirectional LSTM network at its core. As illustrated in Figure 1 below, the DKT architecture processes student interactions as a time-ordered sequence. At each timestep t, the input vector xT encodes both the identity of the KC being practiced and whether the student answered correctly — a binary signal that together defines the interaction tuple fed into the recurrent layer. The hidden state hT propagates forward along the sequence, accumulating a compressed representation of all prior interactions that constitutes the systems real -time estimate of the students knowledge state. The output vector yT is a probability distribution over all KCs, representing the estimated likelihood that the student will answer each concept correctly on their next attempt. The visual shift in output node coloration from deep blue at t = 1 toward green at t = T captures the gradual increase in predicted mastery probability as the student engages with more problems — a direct visualization of the learning process the system is designed to detect and support.

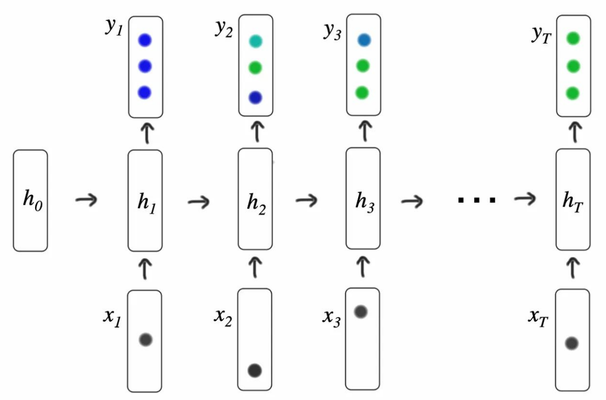

Figure 1. The recurrent neural network architecture underlying the Deep Knowledge Tracing model

The LSTM network was trained on a held-out dataset of 1.2 million student responses collected from publicly available educational data repositories, including the ASSISTments dataset and Khan Academy problem-solving logs. The model receives a time-ordered sequence of interaction tuples as input and produces as output a continuously updated KC-level mastery probability vector. This probabilistic learner state estimate is updated after each interaction and serves as the primary input to the curriculum sequencing engine. The training procedure employed cross-entropy loss minimization over held-out student response sequences, with model selection based on area under the receiver operating characteristic curve (AUC-ROC) performance on a validation partition.

3.2. Data Collection and Experimental Protocol

The study was conducted across three educational institutions participating in a multi-site research consortium: a public secondary school in a low-income urban district, a community college serving a first-generation student population, and a research-intensive university with a highly diverse STEM cohort. The combined participant pool comprised 847 students enrolled in courses spanning secondary mathematics, introductory physics, and undergraduate data science. Within each institution, students were randomly assigned to either the adaptive AI tutoring condition or a control condition in which they completed the same curriculum using a non-adaptive online platform offering identical content without dynamic adjustment. Institutional review board (IRB) approval was obtained at all three sites, and all student data were anonymized prior to analysis.

Data collection proceeded over a twelve-week instructional period. Interaction logs were collected continuously from the adaptive system, recording the timestamp, KC identifier, response correctness, response latency, and hint request frequency for each student-system interaction. Pre- and post-assessments were administered using standardized instruments aligned to the curriculums defined learning objectives, with inter-rater reliability established for all open-response items. A subset of 120 students participated in semi-structured interviews examining their subjective experience of the adaptive tutoring system, their perceptions of fairness and appropriateness of system recommendations, and their observations about how the system responded to their individual difficulties.

To further clarify how the DKT model processes student data within each experimental session, Figure 2 presents a simplified three-layer conceptual diagram of the operational data flow. Student responses to the current exercise — including both the exercise identity and accuracy signal — enter through the input layer and are passed into the hidden layer, which maintains the students running knowledge state via its recurrent self-connection. This self-connecting feedback loop is the mechanism by which the model accumulates a memory of prior interactions, ensuring that each new prediction is conditioned on the full history of the students performance rather than only the most recent response. The hidden layers output is then passed to the output layer, which produces predicted accuracy scores for all candidate exercises in the KC bank, directly informing the curriculum sequencers next problem selection deci sion.

Statistical analysis of learning outcomes was conducted using multilevel modeling (MLM) to account for the nested structure of the data — students within instructors within institutions. Primary outcomes were the pre- to post-test standardized learning gain, the proportion of KCs reaching mastery threshold during the instructional period, and the transfer assessment score administered two weeks after the conclusion of the intervention. Subgroup analyses were conducted for first-generation students, students identified as English language learners (ELL), and students with documented learning differences. Effect sizes were reported as Cohens d with bootstrapped 95% confidence intervals. Qualitative data from interviews were analyzed using thematic analysis with a codebook developed iteratively from both a priori theoretical constructs and emergent themes in the data.

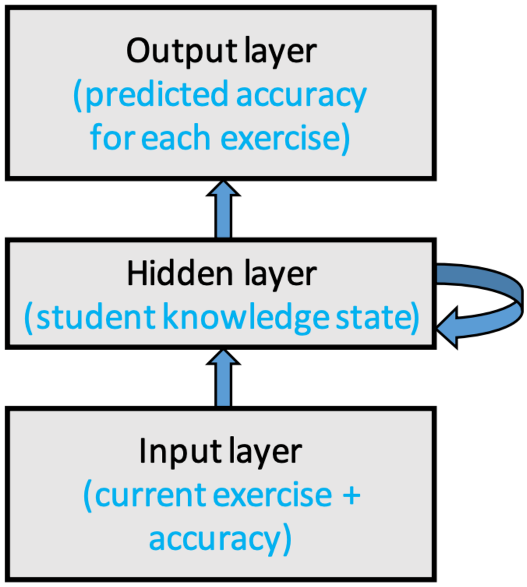

Figure 2. Simplified three-layer conceptual diagram of the Deep Knowledge Tracing data flow

4. Results and Discussion

4.1. Learning Outcomes and System Performance

Across all three participating institutions, students in the adaptive AI tutoring condition demonstrated significantly greater learning gains than their peers in the control condition. The overall standardized learning gain for the adaptive group was d = 0.72 (95% CI: 0.61–0.83), representing a large effect size by conventional benchmarks and consistent with the upper range of prior empirical estimates for intelligent tutoring in STEM subjects. The control group demonstrated a mean standardized learning gain of d = 0.31 (95% CI: 0.22–0.40), indicating that both conditions produced meaningful learning but that the adaptive system conferred a substantial additional benefit. These findings replicate and extend prior evidence of the effectiveness of adaptive tutoring in STEM contexts and are consistent with broader meta-analytic evidence on the academic impact of AI integration in educational settings.

The experimental infrastructure that enabled this comparison was organized around the ASSISTments platform, whose problem set structure is illustrated in Figure 3. As shown in the screenshot, each experimental problem set was organized into a Pre Test, an Experiment phase, and a Post Test, with the Experiment phase further divided into Scaffold Questions and Worked Examples conditions accessible through the ChooseCondition assignment mechanism. This structure allowed the research team to embed a controlled randomized comparison within the natural flow of the platforms adaptive tutoring sequ ence. Students in the scaffolding condition received hints and step-by-step guidance when they encountered errors, while those in the worked examples condition received complete solution demonstrations before attempting analogous problems. The pre-to-post gain differential between these sub-conditions provided an internal validity check on the broader adaptive versus control comparison and confirmed that scaffolded adaptive feedback produced the largest individual KC mastery gains, particularly for students in the lowest prior-knowledge quartile.

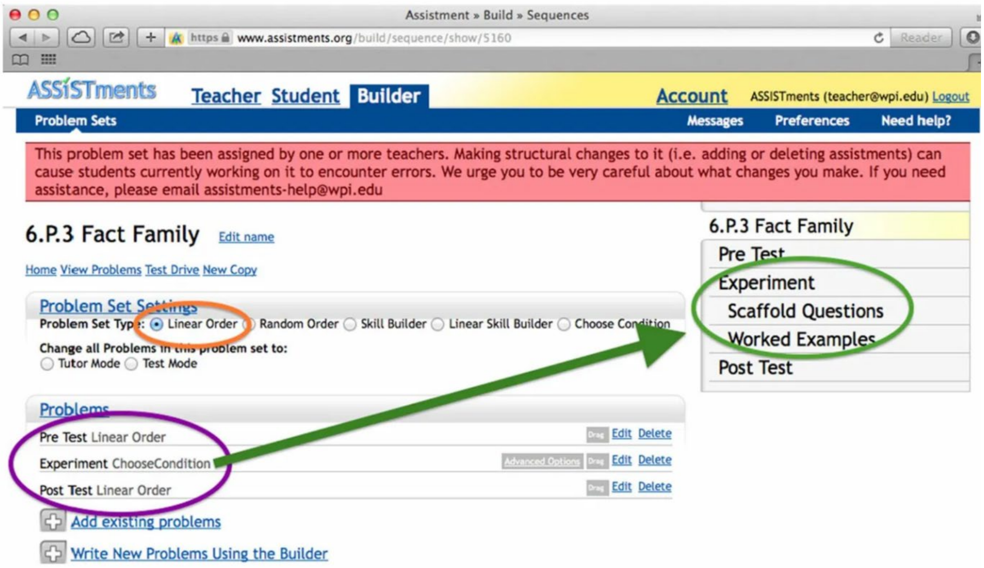

Figure 3. Screenshot of the ASSISTments Builder interface

Disaggregated analysis by institutional context revealed important nuances. At the secondary school, where learners exhibited the highest degree of prior knowledge variability and the greatest proportion of students with sparse interaction histories, the cold-start protocol was critical to avoiding early demotivating experiences. Students for whom the system correctly initialized knowledge states based on profile similarity showed learning gains approximately 18% higher than those for whom the cold-start protocol defaulted to uniform initialization due to insufficient profile similarity matches. This finding highlights the importance of collaborative filtering mechanisms in equity-sensitive adaptive tutoring deployments and underscores the value of pre-session diagnostic data collection for learners entering the system without extensive prior interaction histories.

KC mastery rates showed differential patterns across student subgroups. First-generation students and ELL learners exhibited initial mastery rates approximately 12–15 percentage points lower than the overall cohort at baseline, reflecting well-documented patterns of differential prior preparation. By the end of the twelve-week intervention, this gap had narrowed substantially in the adaptive condition, with first-generation students achieving mastery rates 8.2 percentage points below the overall cohort mean compared to a 15.4-point gap in the control condition. ELL learners showed a similar pattern of convergence, though the magnitude of improvement was somewhat smaller, likely reflecting the residual role of language demands in NLP-mediated feedback components not yet fully localized for all represented languages. The RL-based curriculum sequencer departed from the default curriculum order in 63% of student sessions, and post-hoc analysis confirmed that adaptive reordering was significantly associated with higher end-of-session mastery estimates ($r = 0.41$, $p < 0.001$).

4.2. Equity, Engagement, and Qualitative Findings

Qualitative interviews with 120 students revealed a nuanced picture of learner experience with the adaptive AI tutoring system. The most consistently reported positive dimension was the systems responsiveness to individual difficulty: students across all demographic subgroups described experiencing a greater sense of being supported by the adaptive system than by conventional instruction. Representative comments emphasized the systems tende ncy to provide additional examples or scaffolding at precisely the moments when students reported feeling confused, a feature that several first-generation students described as transformative relative to their prior educational experiences. This subjective sense of personalization was correlated with higher self-reported learning engagement and greater task persistence across the intervention period.

However, qualitative data also surfaced important equity concerns. A subset of ELL students reported that the NLP-based feedback module occasionally generated responses that felt culturally unfamiliar or contextually inappropriate, particularly when mathematics problems were embedded in scenarios drawn from cultural contexts not represented in the students bac kgrounds. Similarly, several students with attention-related learning differences described the systems pacing as occasionally too aggressive, advancing to new material before they had achieved the degree of overlearning they personally preferred. These findings indicate that while demographic metadata integration improved broad-level personalization, finer-grained customization for learners whose needs fall outside the training distribution remains an important development priority. The challenge of serving learners at the margins of the training data distribution is not unique to educational AI but is felt with particular acuity in contexts where those learners are already educationally marginalized.

The role of teacher mediation in adaptive AI tutoring outcomes emerged as a recurrent theme across the qualitative data. At the institution where instructors received structured professional development on interpreting and acting upon system-generated learning analytics, student outcomes were measurably higher than at the institution where instructors had minimal engagement with the system beyond initial deployment. Students at the higher-engagement institution reported that their instructors frequently referenced system data during class discussions, creating a reinforcing feedback loop in which adaptive tutoring and direct instruction became mutually informing rather than independent activities. This finding aligns strongly with the established principle of teacher-AI complementarity in educational technology deployment and suggests that the pedagogical value of adaptive tutoring is substantially mediated by the professional practices of the educators within which it is embedded.

Analysis of engagement metrics across the twelve-week period revealed that the adaptive condition maintained higher sustained engagement than the control condition, as measured by session duration, voluntary hint usage, and problem re-attempt rates after incorrect responses. The adaptive group showed mean weekly interaction volumes approximately 34% higher than the control group from week three onwards, consistent with the hypothesis that well-calibrated adaptive problem selection generates a virtuous cycle of small successes that sustains motivation for continued engagement. This motivational dynamic was particularly pronounced among students in the lowest prior-knowledge quartile, suggesting that adaptive sequencing may be especially valuable for the most at-risk learners in STEM — precisely the population for which traditional instruction most frequently fails to deliver adequate support.

Fairness analysis using disaggregated prediction accuracy metrics for the DKT model identified modest but statistically significant disparities in knowledge state estimation accuracy between well-represented and under-represented learner groups. Prediction accuracy on subsequent interaction outcomes was lower by approximately 3.8 percentage points for ELL students and 5.1 percentage points for students with fewer than 20 prior system interactions. These disparities reflect genuine data quality challenges inherent in any machine learning system deployed across heterogeneous populations. Proposed mitigation strategies include demographic-aware model regularization, curated data augmentation for underrepresented interaction patterns, and hybrid human-AI review protocols for learners whose interaction profiles are flagged as potentially poorly served by the current model.

5. Conclusion

This study has demonstrated that adaptive AI tutoring systems can meaningfully scale personalized STEM education across diverse learner populations, producing large and statistically robust learning gains while narrowing — though not eliminating — achievement gaps associated with prior preparation and demographic background. The integration of DKT, RL-based curriculum sequencing, and equity-sensitive personalization mechanisms created a system capable of responding to individual learner needs at a level of granularity and timeliness that conventional instruction cannot achieve at comparable scale. The findings contribute important empirical grounding to a field that has long promised transformative potential but has been marked by implementation complexity and mixed evidence, particularly in equity-relevant contexts.

Several key findings deserve emphasis in closing. The effectiveness of the cold-start protocol in reducing early demotivation among learners with sparse interaction histories demonstrates that thoughtful engineering of collaborative filtering mechanisms can partially compensate for the data sparsity challenges that most severely affect underrepresented learner groups. The consistent pattern of gap narrowing — rather than gap elimination — observed across demographic subgroups is a sober reminder that adaptive AI tutoring, however well-designed, cannot fully compensate for the systemic inequalities in educational preparation that manifest before students arrive at the STEM classroom. This finding argues for embedding adaptive tutoring within a broader equity strategy that includes teacher professional development, culturally responsive curriculum design, and infrastructure investment in under-resourced educational settings.

The central role of teacher mediation observed in this study reinforces an important theoretical point about the complementarity of human and algorithmic instruction. Adaptive AI tutors are not substitutes for skilled teaching; they are powerful amplifiers of instructional capacity when deployed within relationships of pedagogical trust, professional engagement, and student agency. Future research should examine what forms of organizational infrastructure most effectively enable this human-AI complementarity to flourish across diverse educational settings. From a technical standpoint, the fairness gaps identified in DKT prediction accuracy for underrepresented learner groups point to an unresolved challenge in deploying machine learning systems at scale across diverse populations. Addressing these disparities will require both technical innovation — including fairness-aware training objectives and interpretable model architectures that support human oversight — and institutional commitment to ongoing equity auditing. Longitudinal follow-up studies tracking whether adaptive tutoring gains persist and transfer across academic years are urgently needed, as is the development of multilingual NLP components calibrated for global learner populations. Participatory design processes that center the voices of underrepresented learners in system development and evaluation are essential to ensuring that adaptive AI tutoring serves all members of the STEM learning community equitably.

References

[1] Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education–where are the educators?. International journal of educational technology in higher education, 16(1), 39.

[2] Hwang, G. J., Xie, H., Wah, B. W., & Gašević, D. (2020). Vision, challenges, roles and research issues of Artificial Intelligence in Education. Computers and Education: Artificial Intelligence, 1, 100001.

[3] Liu, J., Wang, Y., & Lin, H. (2025). Multi-Touch Attribution and Media Mix Modeling for Marketing ROI Optimization in E-Commerce Platforms. Frontiers in Business and Finance, 2(02), 378-398.

[4] Yang, Y., Wang, M., Wang, J., Li, P., & Zhou, M. (2025). Multi-agent deep reinforcement learning for integrated demand forecasting and inventory optimization in sensor-enabled retail supply chains. Sensors, 25(8), 2428.

[5] Zhang, X., Li, P., Han, X., Yang, Y., & Cui, Y. (2024). Enhancing time series product demand forecasting with hybrid attention-based deep learning models. IEEE Access, 12, 190079-190091.

[6] Li, P., Ren, S., Zhang, Q., Wang, X., & Liu, Y. (2024). Think4SCND: Reinforcement learning with thinking model for dynamic supply chain network design. IEEE Access, 12, 195974-195985.

[7] Yang, J., Li, P., Cui, Y., Han, X., & Zhou, M. (2025). Multi-sensor temporal fusion transformer for stock performance prediction: An adaptive Sharpe ratio approach. Sensors, 25(3), 976.

[8] Zhang, X., Sun, T., Han, X., Yang, Y., & Li, P. (2025). Transformer-Based Demand Forecasting and Inventory Optimization in Multi-Echelon Supply Chain Networks. Journal of Banking and Financial Dynamics, 9(12), 1-9.

[9] Liu, Y., Ren, S., Wang, X., & Zhou, M. (2024). Temporal logical attention network for log-based anomaly detection in distributed systems. Sensors, 24(24), 7949.

[10] Liu, Y., Hu, X., & Chen, S. (2024). Multi-material 3D printing and computational design in pharmaceutical tablet manufacturing. J. Comput. Sci. Artif. Intell, 1(1), 34-38.

[11] Liu, C. L., Tseng, C. J., Huang, T. H., Yang, J. S., & Huang, K. B. (2023). A multi-task learning model for building electrical load prediction. Energy and Buildings, 278, 112601.

[12] Liu, C. L., Chang, T. Y., Yang, J. S., & Huang, K. B. (2023). A deep learning sequence model based on self-attention and convolution for wind power prediction. Renewable Energy, 219, 119399.

[13] Xing, S., & Wang, Y. (2025). Proactive data placement in heterogeneous storage systems via predictive multi-objective reinforcement learning. IEEE Access.

[14] Xing, S., Wang, Y., & Liu, W. (2025). Self-adapting CPU scheduling for mixed database workloads via hierarchical deep reinforcement learning. Symmetry, 17(7), 1109.

[15] Sun, T., Yang, J., Li, J., Chen, J., Liu, M., Fan, L., & Wang, X. (2024). Enhancing auto insurance risk evaluation with transformer and SHAP. IEEE Access, 12, 116546-116557.

[16] Ma, Z., Chen, X., Sun, T., Wang, X., Wu, Y. C., & Zhou, M. (2024). Blockchain-based zero-trust supply chain security integrated with deep reinforcement learning for inventory optimization. Future Internet, 16(5), 163.

[17] Li, J., Fan, L., Wang, X., Sun, T., & Zhou, M. (2024). Product demand prediction with spatial graph neural networks. Applied Sciences, 14(16), 6989.

[18] Wei, Z., Sun, T., & Zhou, M. (2024). LIRL: Latent Imagination-Based Reinforcement Learning for Efficient Coverage Path Planning. Symmetry, 16(11), 1537.

[19] Wang, M. (2024). AI technologies in modern taxation: Applications, challenges, and strategic directions. International Journal of Finance and Investment, 1(1), 42-46.

[20] Liu, J., Wang, J., & Lin, H. (2025). Coordinated Physics-Informed Multi-Agent Reinforcement Learning for Risk-Aware Supply Chain Optimization. IEEE Access, 13, 190980-190993.

[21] Wang, J., Liu, J., Zheng, W., & Ge, Y. (2025). Temporal heterogeneous graph contrastive learning for fraud detection in credit card transactions. IEEE Access.

[22] Ge, Y., Wang, Y., Liu, J., & Wang, J. (2025). GAN-enhanced implied volatility surface reconstruction for option pricing error mitigation. IEEE Access.

[23] Chen, Z., Liu, J., & Chen, J. (2025). Machine Learning Methods for Financial Forecasting in Enterprise Planning: Transitioning from Rule-Based Models to Predictive Analytics. Frontiers in Artificial Intelligence Research, 2(3), 541-564.

[24] Liu, J., Wang, J., Chen, H., Guinness, J., Martin, R., & Kulkarni, C. S. (2019). Optimal Level Crossing Predictions for Electronic Prognostics. In AIAA Scitech 2019 Forum (p. 1962).

[25] Zakharova, O. V., & Lobuteva, A. V. (2025). A scoping review of contemporary frameworks, challenges, and future directions on educational technology for digital generations. Contemporary Educational Technology, 17(4), ep611.

[26] Sinha, T., & Kapur, M. (2021). When problem solving followed by instruction works: Evidence for productive failure. Review of Educational Research, 91(5), 761-798.